DOCUMENT RESUME

ED 301 165 IR 013 505

AUTHOR Schimmel, Barry J.
TITLE Patterns in Students' Selection of Feedback in Computer-Based Instruction.
PUB DATE Apr 88
NOTE 40p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 4-9, 1988).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; *Anxiety; *College Students; *Computer Assisted Instruction; Correlation; *Feedback; Higher Education; Intermode Differences; Interviews; Multiple Regression Analysis; Pretests Posttests; Rating Scales; Remedial Mathematics; *Student Attitudes

ABSTRACT

This study investigated the relationship between learner attitudes (self-efficacy and anxiety), students' selection of feedback containing different amounts of information (no feedback, correct answer, the explanation alone, and the answer with an explanation), and achievement. Subjects were two clusters of college students (N=49) enrolled in a computerized remedial mathematics course. Data were obtained from pre- and post-instruction measures, online feedback and progress records, and self-reports. Analyses of these data showed that, although individual students exhibited wide differences in selecting feedback information, the choice of feedback followed two general patterns for the two clusters: the 26 students in cluster 1 chose correct response feedback after 71% of their incorrect answers, while the 23 students in cluster 2 showed greater diversity in their selection of feedback information, more frequent selection of feedback information, and more frequent selection of high information feedback. Students in cluster 2 tended to have greater confidence that their answers were correct. It was also found that feedback selection patterns were related to achievement for those students who, after incorrect answers, tended to choose feedback other than the correct response feedback; these students showed better performance, on average, than students choosing other feedback forms. The results of the study suggest at least two possible interpretations regarding cognitive processing: the cluster differences may simply be associated with differences in self-confidence between the students in the two clusters, or the second cluster's diversity in feedback information selection and selection of high information feedback may indicate more intense feedback processing. The text is supplemented by 12 tables, and 41 references are provided. (EW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED301165

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Patterns in Students' Selection of Feedback in Computer-Based Instruction

Barry J. Schimmel

Performance and Learning Associates

Paper presented at the annual meeting of the
American Educational Research Association
New Orleans -- April, 1988

IR013205

© 1988 by
Barry J. Schimmel
Performance and Learning Associates
382 Central Park West 9D
New York, New York 10025

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Barry J. Schimmel

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Patterns in Students' Selection of Feedback in Computer-Based Instruction

Barry J. Schimmel

Performance and Learning Associates

Abstract

The purpose of this study was to investigate student differences in selecting feedback containing different amounts of information. Specific questions were: How do learners differ in their selection of feedback information? Are these differences related to student aptitude? Do these differences affect student achievement? The sample consisted of 49 college students taking a computerized remedial mathematics course. This investigation was a correlational study of learner aptitudes (self-efficacy and anxiety), students' selection of feedback containing different amounts of information (no feedback, the correct answer, and the answer with an explanation), and achievement. Though wide individual differences were found in preference for feedback information, two general patterns of information selection were found.

Patterns in Students' Selection of Feedback in Computer-Based Instruction

Barry J. Schimmel

Performance and Learning Associates

Feedback is a commonly recommended feature of computer-based training (CBT). For example, Cronbach (1977, p.14) described feedback as fundamental to the rationale for CBT, and Gagne, Wager, and Rojas (1984) stated that the inclusion of feedback should be considered a "guideline" in designing CBT. Feedback is also prominent in intelligent tutoring in the form of "coaching" (Goldstein, 1979; Sleeman & Brown, 1982). Coaching has been defined as "supplying advice in those situations wherein the student has chosen a less than ideal option" (Goldstein, 1979). Definitions of feedback have included the following: "information to the learner concerning the correctness" of responses (Frayer & Klausmeier, 1971, p. 19) or "information about the nature and consequences of one's acts" (Cronbach, 1977, p. 89).

In the face of considerable support for the use of feedback in CBT, it is surprising that empirical results on the content of feedback mesages have been so inconsistent. For example, providing learners with extensive information about why their answers are correct or incorrect appears, on the whole, no more useful than offering minimal information (Geis & Chapman, 1971;

This paper is based on the author's dissertation. I am grateful to Professors Lyn Corno of Teachers College, Columbia University and Stephen T. Kerr of the University of Washington for their considerable assistance.

Hendrickson, 1978; Hyman & Tobias, 1981; Kulhavy, 1977).

One explanation for the gap between expectations and empirical findings on feedback content may lie in the manner in which students process feedback. The use of feedback assumes that learners choose to attend to feedback messages, yet little investigation has occurred on the manner in which students process feedback (Kulhavy, 1977). The effects of possible student inattention to certain kinds of feedback have not been distinguished from possible failure of the feedback treatments to increase learning (Winne & Marx, 1980).

The purpose of this study was to investigate student differences in selecting feedback containing different amounts of information. The specific research questions were: (1) How do learners differ in their selection of feedback information? (2) Are these differences related to student aptitude? (3) Do these differences affect student achievement?

Two aptitudes were investigated: self-efficacy and anxiety. Self-efficacy refers to learners' perceived ability to perform a task (Bandura, 1982), similar to self-confidence. Corno and Mandinach (1983) proposed that in comparison with low self-efficacy learners, students with high self-efficacy tend to show more "self-regulation" as learners, more actively selecting relevant from irrelevant information and tending to seek larger amounts of information. Thus, high self-efficacy may be associated with a preference for feedback with considerable information.

Many studies have found that anxiety is associated with difficulty in learning (Benjamin, McKeachie, Lin, & Holinger,

1981; Culler & Hollahan, 1980; O'Neil, Judd, & Hedl, 1977). Wine (1971) hypothesized that part of the attentional capacity of anxious learners is preoccupied with task-irrelevant thoughts such as evaluation concerns, leaving a reduced capacity for task-relevant processing. In particular, she found anxious learners deficient in discriminating relevant and irrelevant cues in problem situations. There is evidence of a negative correlation between anxiety and self-efficacy (Doctor & Altman, 1969; Sarason, 1980). In this study learners with high self-efficacy and low anxiety were expected to seek more information-rich feedback than learners with low self-efficacy and high anxiety.

Students in a computer-based algebra course were able to choose different types of feedback containing varying amounts of information, and their choices, over hundreds of examples, were tracked unobtrusively by the program. Aptitude and achievement tests permitted correlation of the feedback choices with learner characteristics and achievement. In-course interviews with students were conducted to provide additional data to explain students' feedback choices.

METHOD

Treatment

This study was conducted in two phases, with the same treatment in each phase. Students took a computerized version of a required remedial algebra course. The computerized course, entitled "Algebra Mentor," was developed by Professor John C. Miller of The City College of New York, with the feedback modified for this study. The portion of the course included in

this study concentrated on solving signed number examples, such as the following: 3+(-1)=?, 3-8=?, or 4(-3)=?. Verbal problems were not included.

The instructional sequence for performing each algebraic operation was similar. A step-by-step explanation of how to solve an example was presented, illustrated with a number line. Additional examples were then presented for the student to solve, first with a number line, then without. Students progressed by forced mastery; to advance from one lesson to the next, a preprogrammed number of consecutive examples had to be answered correctly. Other program characteristics included backward and forward branching and random generation of examples.

Feedback was available to students after each practice example. The feedback was divided into two student-controlled components. When a student completed an example the following question appeared at the bottom of the screen: "Would you like to see the answer (Y or N)?" If the student responded "yes" and the answer was correct, a beep sounded and "Your answer is correct" appeared onscreen with the example. If the answer was incorrect, the statement "Here is the correct answer:" appeared with the correctly solved example.

After either feedback or the student's rejection of feedback, the question "Do you want an explanation (Y or N)?" appeared at the bottom of the screen. If the student responded positively, an explanation was shown, followed by the question "Shall I repeat the explanation (Y or N)?". A repetition consisted of the same explanation presented again. After the explanation, or if the student did not want the explanation, the

next example or lesson appeared.

Thus, students could receive four types of feedback information: (a) no feedback (by answering "no" to both queries), (b) the correct answer, (c) the explanation alone, or (d) the correct answer and the explanation. These options were available after every example. Students' feedback selections and performance were unobtrusively recorded by the computer.

At five points during the course students were interviewed about the example they were solving at the moment. The interviews were conducted as soon as a student advanced to a lesson in which graphic support (the number line) was not provided. Pilot testing had indicated that errors were most common immediately after the removal of graphic support, and students' feedback choices after incorrect answers were of particular interest due to research suggesting that feedback is more effective after incorrect answers than after correct answers (Kulhavy, 1977).

Instrumentation

Students were given four written pre-instruction measures, two post-instruction measures, and were interviewed. Pre-instruction measures included a test of ability in performing algebraic operations with signed numbers, a measure of self-efficacy, and three measures of anxiety. Post-instruction measures consisted of two tests of achievement and a questionnaire about previous computer experience. In addition to these instruments, two methods were used to record students' processing of feedback, an unobtrusive online record and an

interview protocol. This section will briefly describe all these measures, with more detail available in Schimmel (1986). Reliability data will be presented in the following section.

Pretest. The pretest of ability in signed numbers consisted of 15 items identical in form to examples presented in the course, with the numbers changed. Due to the program's random generation of examples, some pretest examples may have appeared in the program. However, this possibility was not considered likely to affect course performance, given the large number of program examples and the probability that course and pretest examples would be identical in only a few cases, if at all.

Self-efficacy. The self-efficacy measure was similar to that described by Bandura and Schunk (1981) and Schunk (1981) and Salomon (1984). The test consisted of practice items and 5 test cards, all displayed for two seconds each. Each test card showed two examples similar in form to those in the signed numbers course. For each card students were asked "How easy would it be for you to learn how to do these math problems?" Students responded on a test page by circling a number on a number line demarcated in tens from 10-100. Students circled 10 if they were sure they could not learn the examples, and circled higher numbers to indicate increasing expected ease in learning the examples. The brief exposure period allowed students to see types of problems but not to solve the problems. Thus, students' judgments were based on estimates of their capabilities, rather than on concrete evidence. The numbers circled by a student were summed and divided by the number of items. Three practice items preceded the test cards, each containing a pair of foreign words

or phrases. On the same 10-100 number line as the test items, students responded to the question "How easy would it be for you to learn to say these words with the right accent?"

Test Anxiety Scale. Developed by Sarason (1978), the Test Anxiety Scale (TAS) contained 37 statements describing feelings about test-taking. The items were to be marked "true" or "false," with each item counting as one point. The TAS has been used primarily as a measure of test anxiety. Dependent variables correlated with the TAS include both long-term achievement in a course and achievement on tests given in close time proximity with the TAS (Culler & Holahan, 1980; Sarason, 1978; Tobias, personal communication, December, 1984). Sarason (1978) reported test-retest reliability levels over .80 "for intervals of several weeks."

Worry-Emotionality Scale. Used as a second measure of anxiety, the Worry-Emotionality Scale was developed by Morris and Liebert (Morris, Davis, & Hutchings, 1981). "Worry" (W) refers to the cognitive components of anxiety, such as self-preoccupation and persistent concerns about failure. "Emotionality" (E) refers to physiological and emotional components of anxiety, such as tension or increased heart rate (Morris, Kellaway, & Smith, 1978). Typical findings with the scale show that W and E are positively correlated (Deffenbacher, 1980) and that W correlates negatively with "Expectancy," a concept closely related to self-efficacy (Doctor & Altman, 1969; Liebert & Morris, 1967; Morris & Fulmer, 1976). Expectancy was measured by an item at the end of the W-E Scale that asked

students to indicate their expected percentage score on an examination.

The ten item W-E Scale consisted of five W items and five E items intermixed. Students responded to each item on a scale of 1 to 5 in relation to the pretest they had just completed. An illustrative item was "I felt that others would be disappointed in me." These ratings were summed, so that the possible range of scores on each measure was 5-25.

The original research (Morris & Fulmer, 1976; Morris & Liebert, 1970) on the Worry-Emotionality Scale was done with college students and middle class high school seniors of apparently average ability. For this study, with participants of low verbal ability, a practice item with an additional explanation was added to the test.

Interviews. The purpose of the interviews was to determine students' (a) confidence that their answer was correct, (b) knowledge of the procedure for solving the example displayed, (c) reasons for making the feedback choices observed, and (d) rating of feedback helpfulness. Consistent with Ericsson and Simon (1980), the self-reports immediately followed the cognitions reported.

Categories for coding the interview data were developed by this researcher iteratively, as he and an assistant coded the data. The data were then coded independently by two other individuals.

Posttests. Two tests of achievement were used. The first contained 15 items similar to those in the pretest. The second test was a seven-item measure of transfer. The transfer test

included examples requiring both signed number skills and skills taught later in the programmed course, and two examples written in verbal, rather than numerical, form. The criteria for transfer items are imprecise (Royer, 1979), but the items used in this transfer test appeared to be a reasonable sample of the signed number tasks that would face these students after the course.

Sample

Participants were students taking a remedial algebra course at The City College of New York. City College does not require standardized student test data. However, other indicators of low academic ability were available. About 60% of the students were receiving financial aid through a program requiring both low family income and low academic performance (cumulative high school grade average of less than 80, graduation in the lower two thirds of one's high school class, or a combined Scholastic Aptitude Test score of less than 900). The content of the remedial course repeats much of the mathematics taught in earlier school years, yet roughly 3/5 of the class usually fails. About 65% of the group were female. Almost half the group were computer novices, while about one third had over six months' computer experience in word processing, programming, computer games, or CBT.

At the beginning of the first phase, all 142 students attending several sections of the remedial class were tested. Students with pretest scores over 67.7 % were eliminated, on the premise that higher-level students would gain little from a

signed numbers course. Eighty-six students (60.6% of the group) remained. These students were ranked by self-efficacy scores, and 29 students were randomly selected from the upper and lower self-efficacy quartiles. Two of these students did not begin the course, and two others who failed to complete more than 60% of the course were excluded from the sample. Thus, the first phase sample contained 25 participants.

The selection procedure employed in the first phase was to be used in the second phase, and the 46 students present at the beginning of the second course were given all the pre-instruction tests. However, large numbers of student absences in the early days of the second course cycle meant that the selection procedure would yield a very small sample. Therefore, all students present on the day the course was to begin were chosen for the CBT course, without using pretest or self-efficacy scores as selection criteria. Of the 25 students who began the course, one failed to complete it and was dropped from the sample, leaving a final sample of 24. None had participated in the first phase of this study.

Each sample in the study had certain advantages and disadvantages with respect to the tasks presented. The second sample contained a larger proportion of females and low income students than the first sample, groups that often show comparatively low achievement in mathematics (Carpenter, Corbitt, Kepner, Lindquist, & Reys, 1981; Fennema & Carpenter, 1981). However, the second sample had just reviewed the subject matter in the first course cycle, and was not selected for its low pretest scores.

Procedure

The computerized course was part of Mathematics 54, a remedial mathematics course historically given several times per semester. Students failing the course during the first cycle retook it in the second cycle of the semester.

At the initial class of each phase a research associate explained that a new but equivalent version of the course would be offered, and that students for the new version would be chosen on the basis of the "exercises" to be completed during that period. The pre-instruction tests were completed in the following order: self-efficacy, Test Anxiety Scale, pretest, and Worry-Emotionality Scale. Though untimed, the tests were all completed within 35 minutes. Three days later, with no intervening instruction, the selected students convened in a separate room, where the principal researcher explained the CBT course. Students then completed a research consent form and signed up for periods to use the microcomputers.

With the exception of sample selection, the procedures for the two studies were almost identical. The microcomputers were located in a Resource Center containing seven microcomputers. At the first session students received training in the operation of the microcomputer and the program. After training, technical questions or problems were infrequent and almost all were handled by a proctor. Two proctors were available in the Center at all times. They provided technical assistance, administered diskettes, maintained records of student progress, interviewed students, supervised administration of post-instruction measures,

and discouraged students from assisting each other. Proctors did not offer instruction in course content.

Students worked on the program for up to 8 periods weekly. While taking the programmed course, students did not attend their regular mathematics classes. Participants completed the signed numbers course at their own pace, and then received the posttest. At five designated points in the course students were interviewed.

Data Analysis

Several sets of data were obtained from the pre- and post-instruction measures, online feedback and progress records, and self-reports. Quantitative data were analyzed using the Statistical Package for the Social Sciences (Nie, Hull, Jenkins, Steinbrenner, & Brent, 1975). Each data set was summarized, with intra-correlations noted. Correlations among data sets were computed. Feedback selection patterns were analyzed to determine whether patterns existed within individual learners and across learners, and variables associated with group feedback selection patterns were identified. The effect of feedback on achievement was analyzed using multiple regression.

RESULTS

Data were collected on pre-test and aptitude, process, and achievement variables. Given the large number of variables in this study, it will be useful to keep in mind the main research questions: (1) How do learners differ in their selection of feedback information? (2) Are these differences related to student aptitude? (3) Do these differences affect student achievement?

Reliability of Measures

Alpha coefficients were computed for all pre-instruction and post-instruction measures (see Table 1). As expected in a group selected for homogeneous mathematics ability, pretest reliability was below .80 in all administrations, and generally declined as sample size and heterogeneity decreased from first to second sample. Reliability levels of the other pre-instruction measures, ranging from .70 - .90, were similar to those found in previous studies with these measures (e.g. Morris & Liebert, 1970; Salomon, 1984; Sarason, 1978; Schunk, personal communication, December, 1984).

Reliability levels for post-instruction measures also decreased from the first to the second sample as group heterogeneity decreased. The alpha coefficient on the different items posttest was particularly low (.25). To increase overall posttest reliability, the tests of similar and different items were combined. The effect was to increase the posttest reliability level to .72 for the total group. Inter-coder agreement on the categorization of open-ended self-report interview items varied from 71.2 - 89.5%, on all items but one, which showed a reliability of 50.0%.

To determine whether students' self-reports were influenced by student or interviewer gender, two-way analyses of variance were conducted. Student and interviewer gender were used as independent variables in analyses of variance with learners' response confidence and other self-report results. No significant F values were found either for the main effects of student or interviewer gender, or the interaction.

In sum, the reliability of the pretest and combined posttest were .74 and .72 respectively for the combined group. The reliability of the other pre-instruction measures was consistent with previous uses of the measures. The coding of self-report data appeared reliable (with inter-coder agreement on all but one item ranging from roughly 71% - 90%) on all but one coding category.

Pre-Instruction Data

A one way analysis of variance was used to test for pre-instruction differences within the samples. Significant differences in pre-instruction results associated with gender, membership in the financial assistance program, and previous computer experience were not found. The first and second samples were compared using a one way analysis of variance on the pre-instruction scores. The samples were found to differ significantly only on pretest scores ($F=28.02$, $p = .001$), with the second sample showing higher scores. Since no between-sample difference was found on the other pre-instruction measures, the two samples were pooled to provide a larger sample for statistical analyses. Further results will be based on the pooled sample, but will take into account the important pretest differences between samples.

Pre-instruction data from the first, second, and combined samples are summarized in Table 2. The pretest mean increased from the first sample to the second. This increase was expected since low pretest scores were a deliberate selection criterion in the first sample. The maximum self-efficacy score was attained

in both samples. Test Anxiety Scale means and standard deviations were similar to those reported by Sarason (1978).

Correlations among the pre-instruction measures for the combined sample are shown in Table 3. As expected, the three anxiety measures were positively correlated. Little correlation was found between self-efficacy and anxiety for the combined sample, probably due to the low variance found among the self-efficacy scores of the second sample.

General Study Characteristics

Participants' study characteristics, recorded online and in observations made regularly by proctors, varied widely (see Table 4). Course completion time varied from 1 - 7 sessions or from 1.4 to 11.6 hours. The mean percentage of examples answered correctly was 87.1%, indicating that students generally found the program easy. The final entry in Table 4 shows the amount of off-line study time reported by learners.

The typical learner spent 3 - 5 sessions, or 4.7 hours, in the course, and completed virtually all 57 lessons, including slightly over 300 examples with about 87% correct. Three students did not complete all lessons, missing the last 1-3 lessons each. Study time outside the computer room was reported as brief: a mean .9 hours, or about 12.5 minutes per session.

Feedback Selection Patterns

After each example learners could choose among four types of feedback information: (1) no feedback, in which students selected neither the correct answer nor an explanation of how to find the answer; (2) "yes-no" feedback (or correct response feedback), where students chose to see the correct answer but no

explanation; (3) "no-yes" feedback, in which learners did not want to view the correct answer but chose to see an explanation; and (4) full feedback (correct response and explanatory feedback, or high information feedback), in which learners saw both the correct answer and an explanation.

Table 5 indicates the feedback patterns recorded. Feedback patterns are shown as the percentage of each student's examples in which the learner selected each type of feedback. The selection of no-yes feedback was negligible. Therefore, to simplify data presentation, this type of feedback will not be shown in further analyses.

Feedback patterns varied widely. For example, at least one student never selected no feedback, while another student chose it on two out of three examples. A third student used full feedback 25 times more than another student.

Table 6 shows the percentages of the three main feedback patterns after correct and incorrect examples separately. (Mean percentages within each section do not sum to 100% due to rounding and the exclusion of no-yes feedback.) The most common feedback pattern after both correct and incorrect answers was yes-no, or correct response, feedback. Learners selected more feedback information after incorrect examples than after correct examples. Explanations were chosen after only 3.1% of correct examples, but after 35.5% of incorrect examples. Similarly, students chose no feedback more frequently after correct examples than after incorrect answers.

Many students tended to choose the same type of feedback

information repeatedly. Table 7 shows two definitions of a "dominant feedback pattern": students choosing a certain type of feedback after more than 70% or 50% of their examples. The upper section indicates the numbers of learners choosing a particular type of feedback information in more than 70% of their feedback decisions. Of the total sample, 79.6% chose yes-no feedback after more than 70% of their correct answers; 28.6% selected this feedback after more than 70% of their incorrect examples; and 18.4% of students showed no dominant pattern (NDP) after correct examples. A similar analysis is shown using the 50% criterion for dominant feedback patterns. Students exhibited greater diversity in feedback patterns after incorrect answers than after correct answers. This is apparent from the last column, which indicates that "no dominant pattern" was more common after incorrect examples.

Correlations among students' use of the six types of feedback information are shown in Table 8. After correct examples, no feedback and yes-no feedback showed an almost perfect negative correlation (-.95), indicating that a learner choosing one of these feedback types during the course almost never wanted the other type. A high negative relationship appeared after incorrect examples between Yes-No and full information feedback. These high negative correlations indicate that, after correct examples, no feedback and yes-no feedback were chosen by largely separate groups of learners. After incorrect answers, primarily separate groups chose yes-no and full information feedback.

To identify additional patterns in students' feedback selections, a cluster analysis was done on the six types of

feedback, using the BMDP Statistical Software package (Dixon, 1981, p. 360). Two clusters of learners were identified, containing 26 and 23 students respectively. Each cluster's use of the types of feedback information is shown in Table 9. After incorrect answers students in cluster 1 selected yes-no feedback after 71.8% of their examples, whereas those in cluster 2 showed more variation in their feedback choices (no feedback 10%, yes-no feedback 41%, and full information feedback 49%). Cluster 2 displayed a more diverse pattern of feedback selection and greater preference for full information feedback than cluster 1.

To summarize, feedback patterns were represented as the percentage of a learner's feedback for which that learner wanted a given type of feedback. In general, three types of feedback were selected: no feedback, correct response (yes-no) feedback, and correct response with an explanation (full information). Learners chose more feedback information after incorrect answers than after correct answers, although correct response feedback was the most frequent choice. Many learners exhibited substantial consistency in their patterns of feedback selection, choosing a particular type of feedback for more than 50% or even 70% of their feedback. Feedback selection was more varied after errors than after correct answers. A cluster analysis identified two overall patterns in selecting feedback information. One cluster chose mostly yes-no feedback. The other cluster showed greater diversity in feedback selection, and selected full information feedback more frequently.. These patterns were particularly interesting, and formed the basis for much of the

later data analysis.

Self-Report Data

At 5 points in the CBT course students were interviewed as they worked on an example. The mean interview time was 4.0 minutes (range 2.0 - 5.6 minutes). Among other items, students were asked about their level of confidence that their answer to the example was correct. The mean confidence level was 3.61 measured on a 4-point scale (4 is high) before feedback (SD=.35, range 2.6 -4.0). The interviews also produced data on students' reasons for the feedback selections they made and on students' understanding of the algebra operations in the program. These self-report data had little effect on the results of the study and will not be reported here.

Post-Instruction Data

The posttest (combined test of similar and transfer items) results for the first sample, second sample, and combined sample are shown in Table 10. One student reached the maximum possible test score of 22, and 26.5% of the sample scored over 90%. A logarithmic transformation of the posttest data was performed to yield a more normal distribution of the posttest data. The results were almost identical to those found with the original scores. Consequently, only the original posttest scores will be used here.

A t test for independent samples was used to test for differences between the first and second sample on the posttest. No significant difference were found, further supporting the appropriateness of aggregating the samples.

Correlates of Feedback Patterns

A summary of the demographic, pre-instruction, and process variables that correlated significantly with feedback use patterns is presented in Table 11. In the last column the table presents point-biserial correlations between these variables and the cluster analysis groups described earlier. The point-biserial correlations indicate that inclusion in cluster 2, characterized by more diverse use of feedback and more frequent use of full information feedback, was positively correlated with pretest, response confidence, percentage of correct answers, and sample (i.e. was more likely to occur in the second sample). Cluster 1 consisted largely (65.4%) of sample 1 learners, whereas cluster 2 contained mostly (65.2%) sample 2 students. (Although the second sample included more females than did the first sample [75% vs. 56%], the difference was not large enough to establish a significant correlation between gender and the feedback clustering.

To identify learner characteristics associated with the two major patterns of feedback use (represented by each cluster), a discriminant analysis of the clusters was done. The variables selected for the analysis were those on which t tests showed significant differences between the two feedback clusters (using two-tailed t tests for independent groups) and which were correlated with each other. Three variables passed this test and were entered into a stepwise discriminant analysis: pretest, response confidence, and sample. With two groups to be discriminated, the analysis could identify only one discriminant function (Nie, Hull, Jenkins, Steinbrenner, & Brent, 1975). This

function consisted of response confidence and sample. A chi square of 6.957 ($p = .031$) indicated that the function's ability to discriminate between the two clusters was significantly better than chance. If used to predict students' membership in their feedback selection clusters, the function would correctly classify 73.5% of the students (80.8% of those in cluster 1 and 65.2% of those in cluster 2).

Correlates of Achievement

Variables significantly correlated with achievement are presented in Table 12 in the order of their correlation with the posttest. Note that sample is not included in Table 12; though correlated with feedback use, it was not strongly associated with posttest results. The strong negative correlations between achievement and study time both in and out of the course may appear surprising, in light of research demonstrating a positive correlation between achievement and study time (e. g. Denham & Lieberman, 1980). In this study, however, in-course study time is an artifact of the forced mastery feature of the computer program used. The more examples a student answered incorrectly, the more examples were presented. Reported out-of-course study time appears to reflect student anxiety; learners studied when they were anxious about their progress, as indicated by the negative correlations between outside study time on the one hand, and self-efficacy and students' confidence during the course on the other ($r = -.26$ and $-.33$ respectively). Outside study time was also positively correlated with anxiety ($r = .26$).

To determine the relative influence of these variables on

achievement, posttest scores were regressed on the achievement correlates. Due to the relatively small sample size, a stepwise multiple regression was used that selected only two variables, those with the highest correlations with the posttest: pretest and outside study time. As expected, pretest explained almost 35% of the variance in achievement. Outside study time accounted for about 5% of the variance.

To "scan" a wider range of variables that might predict achievement, a second stepwise multiple regression analysis was done for which variable selection was not limited by the bivariate correlations with achievement. The many intercorrelations among the variables in Table 12 necessitated either combining or eliminating variables. Pretest and self-efficacy, with a correlation of .38, were combined as "aptitude" for learning. Both aptitudes have been shown to have significant effects on learning (e.g. Bandura & Schunk, 1981; Snow & Lohman, 1984). The four predictors entered into the stepwise regression were aptitude, gender, yes-no feedback after incorrect answers, and yes-no feedback after correct answers. The highest bivariate correlation in this group was .28. The regression equation was set to accept two predictors. Table 13 presents the results. The two predictors identified in the regression equation were aptitude, as expected, and yes-no feedback after incorrect answers. These predictors accounted for about 44% of the variance in achievement. Feedback selection (i.e. yes-no feedback after incorrect answers) predicted about 7% of the total variance.

To summarize, two stepwise multiple regression analyses were

completed to determine the predictors of achievement. Pretest and outside study time reported by students were identified in the first analysis as chief predictors of feedback patterns. Pretest and self-efficacy were then combined to form a general aptitude for learning. A second multiple regression analysis identified learning aptitude and yes-no feedback after incorrect answers. These variables explained 37% and 7% of the posttest variance respectively.

Feedback selection cluster 2, with greater diversity in their selection of feedback information and greater use of full information feedback, scored significantly higher on the posttest than cluster 1. This learning advantage was achieved independently of pretest differences.

DISCUSSION AND CONCLUSIONS

The key findings in this study were the following: (1) Individuals exhibited wide differences in selecting feedback information. For example, one student chose feedback containing the correct answer and an explanation 25 times more than another student. (2) Two general patterns were found in students' selection of feedback information. In the first pattern, exhibited by 26 learners (cluster 1), students chose correct response feedback after 71% of their incorrect answers. In the second pattern, shown by 23 students (cluster 2), students showed greater diversity in their selection of feedback information and more frequent selection of high-information feedback. (3) Learners in cluster 2 tended to have greater confidence that their answers were correct, and tended to be in the second

sample. (4) Feedback selection patterns were related to achievement; students who, after incorrect answers, tended to choose feedback other than correct response feedback showed better performance, on average, than students choosing the other feedback forms.

Although pretest and sample were significantly correlated ($r = .61$), only sample was identified as a predictor of feedback clustering (in combination with response confidence or percentage correct). Cluster 1 contained mostly sample 1 learners, whereas cluster 2 consisted largely of sample 2 students. Clearly, one or more unidentified sampling factors exerted considerable influence on the selection of feedback information.

A surprising finding was that feedback "cluster" was not as good a predictor of achievement as percentage of yes-no (correct response) feedback after incorrect answers (the less correct response feedback selected after incorrect answers, the better the learners' achievement). One might think that a variable encompassing all three types of feedback would predict achievement better than a single feedback type. However, when "cluster" was substituted for individual types of feedback in the second multiple regression, in which aptitude, gender, and two types of yes-no feedback were entered, feedback was not selected as a predictor. The predictors selected in that regression were aptitude and gender. Apparently, it was not the entire feedback cluster that influenced achievement, but the use (or avoidance) of one specific type of feedback: yes-no feedback after incorrect answers.

Cluster 1 was approximately evenly divided by gender;

cluster 2 contained a large proportion (73.9%) of females, although not enough to create a statistically significant gender difference betwen the groups. The preponderance of females in cluster 2 may have inclined that group toward the greater use of feedback observed, as suggested by a Peterson and Fennema study (1985). They found that the mathematics achievement of fourth grade females was positively correlated with the amount of assistance learners obtained, such as teacher assistance or group cooperation. Males exhibited the reverse pattern. If feedback operates as a kind of assistance, feedback may be more frequently sought and more effectively used by females than males. The achievement variance predicted by cluster as a whole appears largely gender-related.

The sample size in this study did not permit the separation of pretest and sampling effects on achievement. The high correlation between sample and pretest suggests that both variables influenced achievement. However, with the relatively small sample used, only one of these variables could be entered into the multiple regression without risking multicollinearity between variables.

The results of this study suggest at least 2 interpretations regarding the cognitive processing represented by the feedback selection patterns found. First, taken at face value, the cluster differences may simply be associated with differences in self-confidence between students in each cluster. Alternatively, the second cluster's diversity in feedback information selection and selection of high-information feedback may indicate more

intense feedback processing. Tobias (1984) also viewed students' use of multiple CBT study options as an indicator of depth of processing. He found that learners who used multiple study options showed greater achievement than those who used few options. The correlation in this study between cluster 2 membership and achievement suggests that the cluster 2 pattern reflects deeper processing. This view also fits Salomon's (1984) AIME hypothesis that students who invest more "mental effort" in learning (perhaps by actively selecting information and seeking more feedback information) will show greater achievement.

The patterns of feedback use observed here suggest that some learners were more active users of feedback and wanted more feedback information than others. This active use of information for learning is characteristic of what Corno and Mandinach (1983) termed self-regulated learning. This study suggests that self-regulated learning, which has been associated with high ability (Brown, 1978; Gray, 1983; Mandinach, 1984), is a construct different from ability. In this study, even among learners of uniformly low ability, significant diferences were found in key aspects of self-regulated learning.

References

Bandura, A. (1982). Self-efficacy in human agency. American Psychologist, 37, 122-147.

Bandura, A. and Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. Journal of Personality and Social Psychology, 41, 586-598.

Benjamin, M., McKeachie, W. J., Lin,Y., and Holinger, D. P. (1981). Test anxiety: Deficits in information processing. Journal of Educational Psychology, 73, 816-824.

Brown, A. L. (1978) Knowing when, where, and how to remember: A problem in metacognition. In R. Glaser (Ed.), Advances in instructional psychology

Carpenter, T. P., Corbitt, M. K., Kepner, H. S. Jr., Lindquist, M. M., and Reys, R. E. (1981). Results from the second mathematics assessment of the National Assessment of Educational Progress. Reston, VA: National Council of Teachers of Mathematics.

Corno, L. and Mandinach, E. B. (1983). The role of cognitive engagement in classroom learning and motivation. Educational Psychologist, 18 (2), 88-108.

Cronbach, L. J. (1977). Educational psychology. NY: Harcourt Brace Jovanovich.

Culler, R. E. and Holahan, C. J. (1980). Test anxiety and academic performance: The effects of study-related behaviors. Journal of Educational Psychology, 72, 16-20.

Deffenbacher, J. L. (1980). Worry and emotionality in test anxiety. In I. G. Sarason (Ed.), Test anxiety: Theory, research, and applications. Hillsdale, NJ: Erlbaum.

Denham, C. and Lieberman, A. (Eds.). (1980). Time to learn. Washington, D.C.: National Institute of Education.

Dixon, W. J. (1981). BMDP Statistical Software 1981. Los Angeles: University of California Press.

Doctor, R. M. and Altman, F. (1969). Worry and emotionality as components of test anxiety: Replication and further data. Psychological Reports, 24, 263-268.

Fennema, E. and Carpenter, T. P. (1981). Sex-related differences in mathematics: Results from National Assessment. The Mathematics Teacher, 74, 554-559.

Frayer, D. A. and Klausmeier, H. J. (1971). Variables in concept learning: Task variables (Theoretical Paper No. 28). Madison, Wisconsin: Wisconsin Research and Development Center for Cognitive Learning, 1971.

Geis, G. L. and Chapman, R. (1971). Knowledge of results and other possible reinforcers in self-instructional systems. Unpublished manuscript. (ERIC Document Reproduction Service No. ED 049464).

Goldstein, I. P. (1979). The genetic graph: A representation for the evolutiuon of procedural knowledge. International Journal of Man Machine Studies, 11, 51-77.

Gray, L. E. (1983). Aptitude constructs, learning processes, and achievement. Unpublished doctoral dissertation, Stanford University.

Hendrickson, J. M. (1978). Error correction in foreign language teaching: Recent theory, research, and practice. The Modern Language Journal, 62, 387-398.

Hyman, C. and Tobias, S. (1981). Feedback and prior achievement. Paper presented at the annual meeting of the Northeastern Educational Research Association, Ellenville, NY.

Kulhavy, R. W. (1977). Feedback in written instruction. Review of Educational Research, 47, 211-232.

Liebert, R. M. and Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. Psychological Reports, 20, 975-978.

Mandinach, E. M. (1984). The role of strategic planning and self-regulation in learning an

Morris, L. W., Davis, M. A., & Hutchings, C. H. (1981). Cognitive and emotional components of anxiety: Literature review and a revised Worry-Emotionality Scale. Journal of Educational Psychology, 73,

541-555.

Morris, L. W. and Fulmer, R. S. (1976). Test anxiety (worry and emotionality) changes during academic testing as a function of feedback and test importance. Journal of Educational Psychology, 68, 817-824.

Morris, L. W., Kellaway, D. S., Smith, D. H. (1978). Anxiety Rating Scale: Predicting anxiety experiences and academic performance in two groups of students. Journal of Educational Psychology, 70, 589-594.

Morris, L. W. and Liebert, R. M. (1970). Relationship of cognitive and emotional components of test anxiety to physiological arousal and academic performance. Journal of Consulting and Clinical Psychology, 35, 332-337.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Brent, D. H. (1975). SPSS. Statistical package for the social sciences. (2nd ed.). NY: McGraw-Hill.

O'Neill, H. F., Judd, W. A., Hedl, J. J. Jr. (1977). State anxiety and performance in computer-based learning environments. In J. Sieber, H. F. O'Neill, and S. Tobias, Anxiety, learning, and instruction. Hillsdale, NJ: Erlbaum.

Peterson, P. L. and Fennema, E. (1985). Effective teaching, student engagement in classroom activities, and sex-related differences in learning mathematics. American Educational Research Journal, 22 (3), 309-335.

Royer, J. M. (1979). Theories of the transfer of learning. Educational Psychologist, 14, 53-69.

Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. Journal of Educational Psychology, 76, 647-658.

Sarason, I. G. (1978). The Test Anxiety Scale: Concept and research. In C. D. Spielberger and I. G. Sarason (Eds.), Stress and anxiety. Vol. 5. New York: Wiley.

Sarason, I. G. (1980). Introduction to the study of test anxiety. In I. G. Sarason (Ed.), Test anxiety: Theory, research, and applications. Hillsdale, NJ: Erlbaum.

Schimmel, B.J. (1983). A meta-analysis of feedback to learners in computerized and programmed instruction. Paper presented at the annual meeting of the American Educational Research Association, Montreal. (ERIC Document 233 708).

Schimmel, B. J. (1986). Feedback use by low-ability college students in computer-based education. Unpublished doctoral dissertation, Teachers College, Columbia University.

Schunk, D. H. (1981). Modeling and attributional effects on children's achievement: A self-efficacy analysis. Journal of Educational Psychology, 73, 93-105.

Sleeman, D. and Brown, J. S. (1982). Introduction: Intelligent tutoring systems. In D. Sleeman and J. S. Brown (Eds.), Intelligent tutoring systems. NY: Academic Press.

Snow, R. E. and Lohman, D. F. (1984). Toward a theory of cognitive aptitude for learning from instruction. Journal of Educational Psychology, 76, 347-376.

Tobias, S. (1984). Macroprocesses, individual differences, and instructional methods. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Wine, J. (1971). Test anxiety and direction of attention. Psychological Bulletin, 76, 92-104.

Winne, P. H. and Marx, R. W. (1980). Matching students' cognitive responses to teaching skills. Journal of Educational Psychology, 72, 257-264.

Table 1

Reliability Coefficients of Pre- and Post-instruction Measures

	First Sample (n=25)	Second Sample (n=24)	Combined Sample (n=49)
Pre-instruction:			
Pretest	.63	.55	.74
Self-efficacy	.94	.81	.90
Test Anxiety Scale	.85	.83	.84
Worry	.59	.81	.70
Emotionality	.83	.78	.81
Post-instruction:			
Similar items	.70	.61	.66
Different items	.35	.07	.25
Combined test	.75	.66	.72

Table 2

Pre-instruction Scores

	First Sample (n = 25)			Second Sample (n = 24)			Total Group (n = 49)		
	M	SD	Range	M	SD	Range	M	SD	Rang
Pretest	7.48	2.76	0- 10	11.13	1.99	7- 14	9.27	3.01	0- 14
Self-efficacy	82.64	21.30	30-100	89.83	13.97	46-100	86.16	18.26	30-10
Test Anxiety Scale	17.92	6.67	6- 28	19.50	7.20	8- 33	18.69	6.90	6- 3
Worry	10.00	3.44	5- 18	8.96	4.56	5- 19	9.49	4.02	5- 1
Emotionality	8.76	4.04	5- 16	7.50	3.70	5- 18	8.14	3.89	5- 1

Note. The scales on the above measures are as follows: pretest (0-15), self-efficacy (0-100), Test Anxiety Scale (0-37), Worry (5-25), Emotionality (5-25).

Table 3

Correlations Among Pre-instruction Measures (n=49)

	Self-efficacy	Test Anxiety Scale	Worry	Emotionality
Pretest	.38*	.10	-.18	-.04
Self-efficacy		-.17	-.34*	-.36*
Test Anxiety Scale			.59*	.66*
Worry				.76*

* $p < .05$

Table 4

General Characteristics of Course and Study Styles

	M	SD	Range
Number of sessions in course	4.2	1.5	1.0- 7.0
Time in course (hours)	4.7	2.2	1.4- 11.6
Number of lessons completed	56.9	0.5	54.0- 57.0
Number of examples presented	306.9	124.3	123.0-605.0
Percentage of correct examples	87.1	5.0	69.1- 95.5
Reported outside time (hours)	0.9	1.6	0.0- 8.5

Table 5

Overall Patterns in Selection of Feedback (Percentages)

	Mean	SD	Range
No	13.9	18.1	0.0-68.7
Yes-no	79.0	17.7	20.7-95.7
No-Yes	0.1	0.3	0.0- 1.8
Full	7.0	5.1	1.4-34.0

Note. No = student wanted neither the correct answer nor an explanation. Yes-No = student wanted the correct answer but no explanation. No-Yes = student did not want the correct answer but wanted an explanation. Full = student wanted both the correct answer and an explanation.

Table 6

Feedback Patterns After Correct and Incorrect Examples (Percentages)

	Mean	SD	Range
After correct examples			
No	14.0	19.2	0.0-72.7
Yes-No	82.2	19.4	20.7-99.4
Full	3.1	4.8	0.0-32.6
After incorrect examples			
No	7.1	10.4	0.0-42.1
Yes-No	57.2	18.3	21.1-85.7
Full	35.5	16.9	12.8-73.3

Table 7

Percentages of Students Showing a Dominant Pattern in Feedback Selection

	No	Yes -No	Full	NDP
Pattern exceeds 70% of examples				
After correct examples	2.0	79.6	0.0	18.4
After incorrect examples	0.0	26.5	4.1	69.4
Pattern exceeds 50% of examples				
After correct examples	10.2	89.8	0.0	0.0
After incorrect examples	0.0	59.2	14.3	26.5

NDP = no dominant pattern, i.e., none of the feedback levels was selected more than the criterion percentage.

Table 8

Correlations In Students' Use of Types of Feedback

	After correct examples		After incorrect examples		
	Yes -No	Full	No	Yes -No	Full
After correct examples					
No	-.95*	-.11	.59*	-.21	-.17
Yes-No		-.12	-.60*	.28*	.11
Full			-.08	-.26*	.32*
After incorrect examples					
No				-.40*	-.21
Yes-No					-.81*
Full					

* $p < .05$

Table 9

Two-Group Cluster Analysis of Learners, Based on Feedback Selections

	Cluster 1		Cluster 2		
	M	SD	M	SD	F
After correct answers					
No	10.44%	13.81%	18.12%	23.60%	1.991
Yes-No	86.74%	14.40%	77.09%	23.09%	3.157
Full	1.69%	1.60%	4.62%	6.51%	4.940
After incorrect answers					
No	4.22%	5.23%	10.34%	13.59%	4.517
Yes-No	71.83%	8.48%	40.64%	10.33%	134.774
Full	23.94%	7.73%	48.54%	14.71%	55.476

Means are the mean percentage of feedback choices in which a learner selected the type of feedback shown.

* $p < .05$

F= Ratio of between-cluster mean square to within-cluster mean square.

Table 10

Posttest Results

	First Sample (n=25)	Second Sample (n=24)	Combined Sample (n=49)
Mean	16.84	17.92	17.37
SD	3.31	2.60	3.01
Range	9-22	9-21	9-22

Table 11

Variables Showing Significant Correlations With Patterns of Feedback Selection ($p \leq .05$)

	After correct answers			After incorrect answers			
	No	Yes -No	Full	No	Yes -No	Full	Clusters
Gender	-.25						
Previous computer time			.40			.29	
Sample							.31
Pretest				.30	-.28		.30
Percentage of correct examples (total program)	.34	-.34			-.39	-.40	.38
Confidence					-.38	.30	.31

Note. These correlations indicate degrees of association with feedback cluster 2.

Table 12

Variables Showing Significant Correlations With Achievement ($p \leq .05$)

	Combined items posttest
Pretest	.59
Time in course	-.52
Reported outside study time	-.51
Response confidence	.49
Quality of procedural explanations	.43
Self-efficacy	.42
Yes-no feedback (after incorrect answers)	-.39
Percentage of correct examples	.33
Gender	-.31
Clusters in feedback selection	.28
Yes-no feedback (after correct answers)	-.28
No feedback (after incorrect answers)	.28

Table 13

Multiple Regression Analysis of Achievement

Predictor	Multiple R	R^2	R^2 Change	b	Standard Error	df	F
Aptitude	.610	.372	.372	.100	2.406	1/47	27.849*
Yes-no feedback after wrong answers	.667	.445	.072	-.452	2.288	2/46	18.407*

* $p < .05$